

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-224722

(43)Date of publication of application : 21.08.1998

(51)Int.Cl.

H04N 5/765

(21)Application number : 09-025494

(71)Applicant : SONY CORP

(22)Date of filing : 07.02.1997

(72)Inventor : MIZUFUJI TARO  
TAKAHASHI TAKAO  
OTA MASASHI  
AKIBA TOSHIYA  
MURABAYASHI NOBORU  
ARAI NAOHISA  
TOMITA MASAMI

## (54) COMMERCIAL SCENE DETECTOR AND ITS DETECTION METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To allow a detector to detect accurately whether a television signal received at present is a commercial program or a proper broadcast program.

SOLUTION: A silence detection circuit 8 detects a silence period of an audio signal SA. A scene change detection circuit 9 detects a scene change point of an image from a video signal SV for the silence period detected by the silence detection circuit 8. A commercial period detection circuit 10 detects whether or not a time interval of scene change points detected by the scene change detection circuit 9 follows a prescribed rule. A tuner 3 detects a multiplex sound mode SSA of a received audio signal.

## CLAIMS

[Claim(s)]

[Claim 1] A commercial sensing device which detects whether the receiving contents of the television broadcasting signal which consists of an image and an audio signal characterized by comprising the following are commercials.

A non-voice interval detection means to detect a non-voice interval of the above-mentioned audio signal.

A scene change detection means to detect a scene change point of a picture from a video signal in a non-voice interval detected by the above-mentioned silent vocal register detection means.

A scene change interval detecting means which detects whether a rule that a time interval of a scene change point detected by the above-mentioned scene change detection means

is constant is followed.

A voice multiplex-mode detection means to detect the voice multiplex mode of the above-mentioned received voice signal.

[Claim 2]The commercial sensing device according to claim 1 detecting whether the above-mentioned scene change interval detecting means is an abbreviated integral multiple whose time interval of the above-mentioned scene change point is 15 seconds.

[Claim 3]Detect a non-voice interval of a received voice signal and a scene change point of a picture is detected from a received video signal in this non-voice interval. A commercial detecting method detecting that the receiving contents of the television broadcasting signal are commercials if a rule that a time interval of this scene change point is constant is followed and it detects that the voice multiplex mode of the above-mentioned received voice signal is a stereo further.

[Claim 4]The commercial detecting method according to claim 3 being an abbreviated integral multiple whose time interval of the above-mentioned scene change point is 15 seconds.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention]This invention is applied for example to image recording playback equipment like a home videotape recorder and relates to an effective commercial sensing device and detecting method.

[0002]

[Description of the Prior Art]Receive television broadcasting and it records with an image recorder like a videotape recorder. When it is reproduced, the commercials inserted in the middle of edited by a program book are flown. The demand of liking to see only this editing is advanced by the viewer and record which reduced the commercial sections or reproduction came to be performed.

[0003]However, the commercial detecting method adopted as the conventional home videotape recorder was simple strictly and it was difficult to distinguish the volume commercials and on program book automatically and in order to carry out record which carried out exact commercial deletion it could not but depend on the user's handicraft.

[0004]The simple commercial detecting method known conventionally is explained below with reference to drawings. Drawing 8 is a block diagram showing an example of the commercial detector circuit 20 which adopts the conventional commercial detecting method and is carried in a videotape recorder etc.

[0005]In this drawing 8 it restores to the RF signal inputted from the antenna 21 which is an RF input terminal with the tuner 22 and an audio broadcast signal is separated from the television broadcasting signal to which it restored. This audio broadcast signal is supplied to the non-voice interval checking circuit 23 and the voice multiplex-mode signal detection circuit 24.

[0006]The non-voice interval checking circuit 23 restores to the inputted above-mentioned audio broadcast signal to a voice baseband signal and detects a non-voice part from

the amplitude level and detects whether the rule that the time interval of the detection is constant is followed.

[0007] Since the commercial section is generally an abbreviated integral multiple for 15 seconds when the time interval at which a non-voice part appears is an integral multiple for 15 seconds, it can be judged that input signals are commercials. Since a non-voice part is generated here also except a start and end point of commercials, in order to avoid that it serves as an integral multiple for 15 seconds accidentally by chance and judges it as commercials, the non-voice detection produced on the conditions of the above-mentioned interval check, i.e. the integral multiple for 15 seconds, judges it as commercials multiple times for example when it repeats twice and it is made to output as a non-voice interval check output, i.e. a commercial detect output.

[0008] On the other hand in the voice multiplex-mode signal detection circuit 24 the present mode of the inputted above-mentioned audio broadcast signal, i.e. a monophonic recording or a stereo and each mode of two languages are encoded and outputted to 2-bit digital data. For example, the time of a monophonic recording is in "00" and a stereo is encoded with "01" and encodes with "10" by two languages.

[0009] This editing mode detector circuit 25 is provided with the delay circuit 26, D flip-flop 27 and the pulse generating circuit 28. This gives this delay of some for about several seconds in inputting the output of the voice multiplex-mode signal detection circuit 24 into the delay circuit 26 and leads it to the input edge of D flip-flop 27.

[0010] On the other hand, this editing mode detector circuit 25 of this inputs the interval check output of the non-voice interval checking circuit 23 into the pulse generating circuit 28. When time observation of this check output is carried out about 2 minutes and 1 time does not have a non-voice interval check output, i.e. a commercial detect output, either the present input signals are not commercials, it judges that it is this editing and a pulse is made to be built. By making it lead this pulse to the clock end of the above-mentioned D flip-flop, the encoding value of the voice multiplex mode of this editing is always held at the output of the above-mentioned D flip-flop.

[0011] Being delayed to the detect output of the voice multiplex mode in the delay circuit 26, it is for avoiding that a voice multiplex-mode detect output becomes unstable when it is expected that change takes place to the voice multiplex mode simultaneously at the time of the commercial detect output used as the time of non-voice detection and a commercial detect output is inputted into the clock end of a D flip-flop.

[0012] One commercial's being less than 1 minute generally and the treatment for shortening response time make time to observe the commercial detect output of the non-voice interval checking circuit 23 about 2 minutes. And the output of above-mentioned D flip-flop 27 turns into an output of this editing mode output circuit 25.

[0013] In the mode comparison circuit 29, the above-mentioned output of this editing mode detector circuit 25 and the output of the voice multiplex-mode signal detection circuit 24 are measured. If in agreement, the present input signal will be judged to be this editing. If inharmonious, it will be judged as commercials and the result will be outputted.

[0014] When the voice multiplex mode shifts to other modes from the mode of this editing and does not shift to the mode of this editing after that again but generally shifts to the 3rd mode, since a program changes and it is expected that it is a point, the mode comparison circuit 29 outputs the decision result of this editing / commercial indeterminate in this case.

[0015]

[Problem(s) to be Solved by the Invention]By the waysince the above-mentioned commercial detector circuit 20 is performing commercial detection only by the speech information of the interval of the voice multiplex mode and a non-audio parterraneous detection will occur frequently considerably.

[0016]In the case of a stereo or bilingual broadcastthe method of preventing erroneous detection had none of especially these editing broadcasts. Although the necessity of carrying out commercial detection from the former like \*\*\*\* based on speech information was knownthe decisive means which is only a simple method also from a detection ratioand improves detection capacity did not exist.

[0017]Then this invention is made in view of the above-mentioned actual conditionand aims at offer of the commercial sensing device and detecting method which can distinguish correctly whether the television broadcasting signal received now is a volume commercials or on program book.

[0018]

[Means for Solving the Problem]In order that a commercial sensing device concerning this invention may solve an aforementioned problemA non-voice interval detection means detects a non-voice interval of a received voice signaland a scene change detection means detects a scene change point of a picture from a received video signal in a non-voice intervalIt detects following a rule that a time interval of a scene change point is constantby a scene change interval detecting meansand a voice multiplex-mode detection means detects further that the voice multiplex mode of the above-mentioned received voice signal is a stereo.

[0019]A commercial detecting method concerning this inventionIn order to solve an aforementioned problemdetect a non-voice interval of a received voice signaland a scene change point of a picture is detected from a received video signal in this non-voice intervalA rule that a time interval of this scene change point is constant is followedand it detects that the voice multiplex mode of the above-mentioned received voice signal is a stereo further.

[0020]In the starting point and an end point of each commercials which can generally specifically be set without between this editingbetween commercials and commercialsand commercialsa non-voice interval for about 0.1 to 2 seconds appears in remarkable high probabilityAnd that a scene change point of a picture is between the non-voice intervaland a thing [ one business time of each commercials / an abbreviated integral multiple for 15 seconds ]And the commercial broadcast section uses that the voice multiplex mode is a stereo modeand when the section when a scene change point at the time of a non-sound appeared in an abbreviated integral multiple for 15 seconds atand was moreover started by an integral multiple for the 15 seconds is a stereophonic broadcastit judges the section to be the commercial section.

[0021]

[Embodiment of the Invention]It explains referring to drawings for the embodiment of the commercial sensing device concerning this inventionand a detecting method hereafter.

[0022]This embodiment is the commercial detector circuit 4 of drawing 1 which detects whether the receiving contents of the television broadcasting signal which consists of an image and an audio signal are commercialsand it has it in the television broadcasting recorder 1.

[0023]In this television broadcasting recorder 1 it gets over with the tuner 3 and the RF signal inputted from the antenna 2 which is an RF input terminal is inputted into the commercial detector circuit 4 respectively after separating into audio signal  $S_A$  voice multiple-signal  $S_{SA}$  and video-signal  $S_V$ . Especially the tuner 3 detects the voice multiplex mode from a voice multiplex pilot signal etc. and inputs the above-mentioned voice multiple-signal  $S_{SA}$  into the commercial detector circuit 4.

[0024]Although this voice multiplex mode has a monophonic recording a stereo and three kinds of two languages it is sent in code which TEREO calls "1" and is called "0" in this embodiment except a stereo.

[0025]The commercial detector circuit 4 is provided with the following.

The non-voice part detector circuit 8 which detects the non-voice interval of the above-mentioned audio signal  $S_A$ .

The scene change detector circuit 9 which detects the scene change point of a picture from video-signal  $S_V$  in the non-voice interval detected by this non-voice part detector circuit 8.

The commercial section detection circuit 10 used as the scene change interval detecting means which detects whether the rule that the time interval of the scene change point detected by this scene change detector circuit 9 is constant is followed.

The above-mentioned tuner 3 used as a voice multiplex-mode detection means to detect voice multiplex-mode  $S_{SA}$  of the above-mentioned received voice signal.

[0026]In the commercial detector circuit 4 commercial detection is performed using audio signal  $S_A$  inputted from the tuner 3 video-signal  $S_V$  and voice multiplex-mode signal  $S_{SA}$  and a commercial decision signal is outputted.

[0027]The flow of the signal in this commercial detector circuit 4 and processing are explained below. As for audio signal  $S_A$  inputted from the tuner 3 dispersion by a predetermined sampling frequency and quantization with a predetermined quantization level, e.g. A/D conversion processing are performed by the A/D conversion circuit 5. In this embodiment quantization of 16 kHz of sampling frequencies and 16 bits of bit length is performed.

[0028]The voice data digitized with A/D converter 5 is inputted into the non-voice part detector circuit 8 in the general purpose processor 7. In this embodiment this general purpose processor 7 was constituted from a computer and has all realized the non-voice part detector circuit 8 the scene change detector circuit 9 and the commercial section detection circuit 10 by software.

[0029]In the non-voice part detector circuit 8 it is asked for the average sound level of the \*\* frame by calculation and the non-voice interval is detected by whether the average level is smaller than a predetermined threshold. The situation of calculation of the above-mentioned average sound level is shown in drawing 2 and the flow of non-voice interval detection processing is shown in drawing 3.

[0030]First the non-voice part detector circuit 8 incorporates the digital sound data outputted from the A/D conversion circuit 5 at Step S1 of drawing 3. Here since it is processing at the rate of 30 frames in 1 second as shown in (A) of drawing 2 16 kHz (sample)/30 frames digital sound data is incorporated. And at Step S2 absolute value-ization of the above-mentioned sample is performed as shown in (B) of drawing 2 and at Step S3 as shown in (C) of drawing 2 an average level is computed.

[0031] Next it is judged whether the above-mentioned average level is smaller than a predetermined silent threshold at step S4. When smaller than a small, namely silent threshold predetermined [above-mentioned] in the above-mentioned average level, here when it is more than a silent threshold conversely predetermined [above-mentioned] again about a silent decision output, an owner sound decision output is taken out with Step S5 at Step S6. It means that taking out a silent decision output had judged with the frame being a silent interval.

[0032] On the other hand, video-signal  $S_V$  outputted from the tuner 3 is inputted into the commercial detector circuit 4 and is inputted into the delay circuit 6, the delay circuit 11 and the scene change detector circuit 9.

[0033] The delay circuit 6 comprises a frame memory of the predetermined frame number of one frame. From this delay circuit 6, the video signal with which only the time corresponding to a predetermined frame number was delayed is inputted into the scene change detector circuit 9.

[0034] In the scene change detector circuit 9, the output from the non-voice part detector circuit 8, the delayed video signal and the through video signal by which direct supply was carried out without delay from the tuner 3 are considered as an input, and scene change detection in a non-voice interval is performed. It explains referring to drawing 4 for flowing into processing in this scene change detector circuit 9.

[0035] First, the output from the non-voice part detector circuit 8 is received at Step S11. By judging whether the above-mentioned output is a silent judging at Step S12 and progressing to Step S15 if it is an owner sound judging, since the frame is not the starting point of commercials and a terminal point, the output that they are not the starting point and the terminal point candidate point of commercials is taken out. When it judges with it being a silent judging at Step S12, it progresses to the judgment routine of the scene change of Step S13 and Step S14.

[0036] The inter-frame-correlation judging performed at Step S13 compares the delay video signal and through video signal which were inputted and computes the correlation. The method of using correlation of the histogram of (1) for example, the signal level of each picture for a mutually related calculation method (2) the method of taking difference about each pixel of each screen and making the integral value a correlation value and (3) -- each screen is divided into two or more fields and how to calculate correlation in each field and to take majority etc. can be considered. Here, the method of (2) is considered to be the simplest is adopted. It explains referring to drawing 5 for this method of (2).

[0037] Image size is also made with  $n \times m$ , the delayed picture image data and through picture image data, and horizontal coordinates are set to  $i$  and it makes [perpendicular direction coordinates]  $S_{ij}$  the data of the coordinates  $(ij)$  of  $D_{ij}$  and through image drawing for the data of the coordinates  $(ij)$  of  $j$  and delay image drawing.

[0038] The correlation value  $E$  of the picture of two sheets is  $E = \sum \sum a_{abc} (D_{ij} - S_{ij})$  when  $abc$  is made into the function which calculates an absolute value. It can come out and express.

[0039] When the correlation value  $E$  between the pictures computed at the above-mentioned step S13 by Step S14 judges with it being larger than a predetermined threshold, it will be small, progresses to the following step and it outputs [there was a scene change, the degree of correlation of the picture of two sheets, that is] it as a candidate point of the starting point and the terminal point of commercials. When the output value of

Step S13 is smaller than a predetermined threshold it progresses to Step S15 and if this frame is not a candidate point of the starting point and the terminal point of commercials it will be outputted.

[0040] Voice multiplex-mode signal  $S_{SA}$  from the output and the tuner 3 of the scene change detector circuit 9 is inputted into the commercial section detection circuit 10. The commercial section detection circuit 10 has a memory area which memorizes the voice multiplex-mode signal from the output and the tuner 3 from the scene change detector circuit 9 by predetermined time. Generally the commercials in television broadcasting are less than 1 minute the longest in many cases and are preparing the memory for 1 minute. e.g. RAM with the capacity of  $120(\text{second}) \times 30(\text{frame}) \times 2(\text{data}) \times 1(\text{bit})$ .

[0041] And within this RAM if the output from the scene change detector circuit 9 is a candidate point of the starting point and the terminal point of commercials as it is shown in drawing 6 if it is not a candidate point about "1" 0 will be memorized to Scene\_Change [Frame].

[0042] Similarly if the voice multiplex-mode signal from the tuner 3 is a stereo the monophonic recording except a stereo that is and when bilingual 0 will be written for "1" in Audio\_Multi [Frame]. Scene\_Change [Frame] in RAM has a field from the present frame to the frame of one quota.

[0043] The example of detection of the actual commercial section in this commercial section detection circuit 10 is shown in drawing 7. A commercial detection result is shown in (C) of the section summarized to (A) of drawing 7 for every scene change the section summarized to (B) of drawing 7 for every voice multiplex mode and drawing 7.

[0044] The voice multiplex mode is a stereo and since the section is moreover continuing by the integral multiple for 15 seconds the sections 1-4 are detected as the commercial section. Although the section 7 is moreover following the sections 8-9 in 15 seconds since the voice multiplex mode is a monophonic recording commercials are not distinguished.

[0045] The commercial section detection circuit 10 outputs "1" when the frame of one quota is judged to be the commercial section and when it was not the commercial section and distinguishes it outputs "0."

[0046] On the other hand audio signal  $S_A$  outputted from the tuner 3 and video-signal  $S_B$  are inputted also into the delay circuit 11 respectively. After modulating a sound and an image since it is necessary to synchronize with the commercial detecting signal outputted from the commercial section detection circuit 10 when recording on a recording medium a synchronization is taken in this delay circuit 11. Since it has a buffer (memory) for 1 minute in the commercial section detection section 10 for commercial detection in the case of this embodiment it is necessary to perform delay processing for 1 minute in the delay circuit 11. Although realizing by a memory is also possible by this embodiment the relation of capacity has realized this using hard disk drive \*\*.

[0047] The output of this delay circuit 11 is inputted into the modulation circuit 12. Here abnormal conditions called eight-to-fourteen modulation are applied to the signal aspect for writing a sound and a video signal in a recording medium. Processing of data compression such as MPEG and JPEG etc. are included here.

[0048] If the output of the modulation circuit 12 is write-in permission it will be supplied to the writing device 14 by the state of the write-in permission switch 13 and will be recorded on the recording medium 15. The sound and video signal which the write-in permission switch 13 was closed when the commercial detecting signal from the

commercial section detection circuit 10 was "1" and were modulated are told to the writing device 14.

[0049]The writing device 14 has electricmagneto opticalphysical or a function recorded on the recording medium 15 with the above combination for the sound and video signal after the abnormal conditions inputted through the write-in permission switch 13.

[0050]The recording medium 15 is a medium which can record and store the data of a compact disk, a magneto-optical disk, a hard disk, etc. which can be written in.

[0051]Thus since the television broadcasting recorder 1 used as this embodiment is provided with the commercial detector circuit 4 which performs exact commercial detection, it can record only this editing program which removed commercials on the recording medium 15.

[0052][Effect of the Invention]The commercial sensing device concerning this invention detects the non-voice interval of a received voice signal by a non-voice interval detection means. A scene change detection means detects the scene change point of a picture from the received video signal in a non-voice interval. Since it detects following the rule that the time interval of a scene change point is constantly the scene change interval detecting means and a voice multiplex-mode detection means detects further that the voice multiplex mode of the above-mentioned received voice signal is a stereo. Commercials are automatically and correctly detectable from the receiving contents of the television broadcasting signal.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1]It is a block diagram of the television broadcasting recorder provided with the commercial detector circuit used as the embodiment of the commercial sensing device concerning this invention and a method.

[Drawing 2]It is a figure for explaining the calculation of the average sound level for every frame which the non-voice part detector circuit which constitutes the above-mentioned commercial detector circuit performs.

[Drawing 3]It is a flow chart which shows the flow of processing of the above-mentioned silent vocal-parts detector circuit.

[Drawing 4]It is a flow chart which shows the flow of processing of the scene change detector circuit which constitutes the above-mentioned commercial detector circuit.

[Drawing 5]It is a figure for explaining the inter-frame-correlation judging performed in the above-mentioned scene change detector circuit.

[Drawing 6]It is a figure for explaining data processing in RAM with which the commercial section detection circuit which constitutes the above-mentioned commercial detector circuit is provided.

[Drawing 7]It is a figure showing the example of detection of the commercial section which the above-mentioned commercial section detection circuit performs.

[Drawing 8]It is a block diagram of the videotape recorder which adopted the conventional commercial detecting method.

[Description of Notations]

1 A television broadcasting recorder and 3 [ A non-voice part detector circuit and 9 / A SHINCHIENJI detector circuit and 10 / Commercial section detection circuit ] A tuner



and 4 A commercial detector circuit and 7 A general purpose processor and 8

---